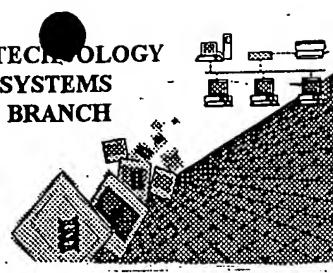BIOTECHNOLOGY
SYSTEMS
BRANCH

# RAW SEQUENCE LISTING
# ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

Application Serial Number:     10/079,137A

Source:     O1PE

Date Processed by STIC:     6 17 02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE
   APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A
   NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
> http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is
a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the
United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission
   User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
   Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
           Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
   2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
   Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

RAW SEQUENCE LISTING                    DATE: 06/17/2002
PATENT APPLICATION:   US/10/079,137A     TIME: 15:42:56

Input Set  :  A:\419c13.app
Output Set:  N:\CRF3\06172002\J079137A.raw

```
 4 <110> APPLICANT: Frudakis, Tony N.
 5         Reed, Steven G.
 6         Smith, John M.
 7         Misher, Lynda E.
 8        ·Dillon, Davin C.
 9         Retter, Marc W.
10         Wang, Aijun
11         Skeiky, Yasir A. W.
12         Harlocker, Susan L.
13         Day, Craig H.
14         Li, Samuel X.
15         Deng, Ta
19 <120> TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
20         AND DIAGNOSIS OF BREAST CANCER
23 <130> FILE REFERENCE: 210121.419C13
25 <140> CURRENT APPLICATION NUMBER: US 10/079,137A
26 <141> CURRENT FILING DATE: 2002-02-20.
28 <160> NUMBER OF SEQ ID NOS: 428
30 <170> SOFTWARE: FastSEQ for Windows Version 4.0
```

**ERRORED SEQUENCES**

```
2636 <210> SEQ ID NO: 158      found 543!
2637 <211> LENGTH: (547) - 
2638 <212> TYPE: DNA
2639 <213> ORGANISM: Homo sapiens
2641 <220> FEATURE:
2642 <221> NAME/KEY: misc_feature
2643 <222> LOCATION: 235
2644 <223> OTHER INFORMATION: n = A,T,C or G
2646 <400> SEQUENCE: 158
2647 cttcaacctc cttcaacctc cttcaacctc ctggattcaa acaatcatcc cacctcagac 60
2648 tccttagtag ctgagactac agactcacgc cactacatct ggctaaattt ttgtagagat 120
2649 agggtttcat catgttgccc tggctggtct caaactcctg acctcaagca atgtgcccac 180
W--> 2650 ctcagcctcc caaagtgctg ggattacagg cataagccac catgcccagt ccatntttaa 240
2651 tctttcctac cacattctta ccacactttc ttttatgttt agatacataa atgcttacca 300
2652 ttatgataca attgcccaca gtattaagac agtaacatgc tgcacaggtt tgtagcctag 360
2653 gaacagtagg caataccaca tagcttaggt gtgtggtaga ctataccatc taggtttgtg 420
E--> 2654 taagtctataccatc taggtttgtg (420) → counted 445!
E--> 2655 taagtaaccat ctaatgatgc atttctcaga 480
E--> 2656 atgtatcctt gtcagtaagc tatgatgtac agggaacact gcccaaggac acagatattg 540
E--> 2657 tacctgt                                                      547
```

```
      3280 <210> SEQ ID NO: 197
      3281 <211> LENGTH: (289) — found 285
      3282 <212> TYPE: DNA
      3283 <213> ORGANISM: Homo sapiens
      3285 <400> SEQUENCE: 197
      3286 ttgggcacct tcaatatcat gacaggtgat gtgataacca agaaggctac taagtgatta 60
      3287 atgggtgggt aatgtataca gagtaggtac actggacaga ggggtaattc atagccaagg 120
E-->  3288 caggaggggtaattc atagccaagg 120 — Counted 145!
E-->  3289 caggactactc aggatagcat gcagtttaaa 180
E-->  3290 acctataagt agtttatttt tggaattttc cacttaatat tttcagactg caggtaacta 240
E-->  3291 aactgtggaa cacaagaaca tagataaggg gagaccacta cgtcgatac         289
      4370 <210> SEQ ID NO: 261
      4371 <211> LENGTH: 746
      4372 <212> TYPE: DNA
      4373 <213> ORGANISM: Homo sapiens
      4375 <220> FEATURE:
      4376 <221> NAME/KEY: misc_feature
W-->  4377 <220> FEATURE:
W-->  4377 <220> FEATURE:
      4378 <221> NAME/KEY: misc_feature
E-->  4379 19, 734, 740, 741              insert <222>
      4380 <223> OTHER INFORMATION: n = A,T,C or G
      4382 <400> SEQUENCE: 261
      4383 ttgggcacct tcaatatcaa tagctaacat ttattgagtg tttatcgtat cataaaacac 60
      4384 tgttctaagc ctttaaacgt actaattcat ttaatgctca taatcacttt agaaggtggg 120
      4385 tactagtatt agtctcattt acagatgcaa catgcaggca cagagaggtt aattaacttg 180
      4386 cccaaggtaa cacagctaag aaatagaaaa aatattgaat ctggaaagtt gggcttctgg 240
      4387 gtaacccaca gagtcttcaa tgagcctggg gcctcactca gtttgctttt acaaagcgaa 300
      4388 tgagtaacat cacttaattc agtgagtagg ccaaatggag gtcagctacg agtttctgct 360
      4389 gttcttgcag tggactgaca gatgtttaca acgtctggcc atcagtwaat ggactgatta 420
      4390 tcattgggaw gtgggtgggc tgaatgttgg ccagtgaagt ttattcawgc catattttta 480
      4391 tgtttaggat gacttttggc tggtcctagg gcaagctctg tctgscacgg aacacagaat 540
      4392 wacacaggga cccccctcaat ttctggtgtg gctagaacca tgaaccactg gttgggggaa 600
      4393 caagcggtca aaacctaagt gcggccggct ggcagggtcc acccatatgg ggaaaactcc 660
W-->  4394 cnacgcgttt ggaatgcctn agctngaatt attctaanag ttgtccncnt aaaattagcc 720
W-->  4395 tgggcgttaa tcangggtcn naagcc                                   746
      4723 <210> SEQ ID NO: 280
      4724 <211> LENGTH: (272) — found 276
      4725 <212> TYPE: DNA
      4726 <213> ORGANISM: Homo sapiens
      4728 <400> SEQUENCE: 280
E-->  4729
( tacctgacat ggagaaataa cttgtagtat tttgcgtgc ggagaaataa cttgtagtat tttgcgtgcatgaat gaactagcaa tgc     insert hard return
E-->  4730 gttgaatgga aaaggtgagt ttcagaagga tatatatgcc ctctaaatcc atttatgtaa 180
E-->  4731 acctttaaaa aactacatta tttatggtca taagtccatc cagaaaatat ttaaaaacct 240
E-->  4732 acatgggatt gataactact gatgtcaggt ca                           272
```

RAW SEQUENCE LISTING ERROR SUMMARY          DATE: 06/17/2002
PATENT APPLICATION:   US/10/079,137A          TIME: 15:42:59

Input Set : A:\419c13.app
Output Set: N:\CRF3\06172002\J079137A.raw

**Invalid Line Length:**
The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:280; Line(s) 4729

VERIFICATION SUMMARY                    DATE: 06/17/2002
PATENT APPLICATION:  US/10/079,137A     TIME: 15:42:59

Input Set : A:\419c13.app
Output Set: N:\CRF3\06172002\J079137A.raw

```
L:95  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:660
L:96  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:720
L:97  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:780
L:98  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:840
L:99  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:900
L:100 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:960
L:101 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:1020
L:122 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:4
L:132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:540
L:133 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:600
L:134 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:660
L:135 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:720
L:136 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:780
L:137 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:840
L:138 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:900
L:139 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:960
L:140 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:1020
L:141 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:1080
L:162 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:5
L:168 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:300
L:169 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:360
L:170 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:420
L:171 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:480
L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:540
L:173 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:600
L:174 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:660
L:175 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:720
L:176 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:780
L:177 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:840
L:178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:900
L:179 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:960
L:201 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:208 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:215 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:222 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:229 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:235 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:239 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:180
L:240 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:240
L:241 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:300
L:242 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:360
L:243 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:420
L:244 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:480
L:245 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:540
L:246 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:600
L:247 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:660
L:248 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:720
L:249 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:780
```

VERIFICATION SUMMARY                      DATE: 06/17/2002
PATENT APPLICATION:  US/10/079,137A       TIME: 15:42:59

              Input Set : A:\419c13.app
              Output Set: N:\CRF3\06172002\J079137A.raw

```
L:250 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:840
L:251 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:900
L:274 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:480
L:277 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:660
L:278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:720
L:279 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:780
L:280 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:840
L:281 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:900
L:282 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:960
L:283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:1020
L:306 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:8
L:345 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:9
L:2110 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:2114 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:132
L:2170 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:2174 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:136
L:2654 M:254 E: No. of Bases conflict, LENGTH:Input:420 Counted:445 SEQ:158
M:254 Repeated in SeqNo=158
L:2657 M:252 E: No. of Seq. differs, <211> LENGTH:Input:547 Found:543 SEQ:158
L:3288 M:254 E: No. of Bases conflict, LENGTH:Input:120 Counted:145 SEQ:197
M:254 Repeated in SeqNo=197
L:3291 M:252 E: No. of Seq. differs, <211> LENGTH:Input:289 Found:285 SEQ:197
L:4377 M:283 W: Missing Blank Line separator, <220> field identifier
L:4377 M:256 W: Invalid Numeric Header Field, <220> has non-blank data
L:4379 M:250 E: Invalid Numeric Identifier, INVALID IDENTIFIER
L:4394 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:261
L:4395 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:261
L:4635 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:275
L:4639 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:275
L:4729 M:254 E: No. of Bases conflict, LENGTH:Input:120 Counted:124 SEQ:280
M:254 Repeated in SeqNo=280
L:4732 M:252 E: No. of Seq. differs, <211> LENGTH:Input:272 Found:276 SEQ:280
L:4783 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:283
```